

Interactions of hadrons in the CALICE SiW ECAL prototype

Philippe Doublet, Michele Fauci-Giannelli, Roman Pöschl,
François Richard for the CALICE Collaboration

This note contains preliminary CALICE results, and is for the use of members of the CALICE Collaboration and others to whom permission has been given.

Abstract

This article presents results of test beams obtained for pions with energies between 2 and 10 GeV which interact in the volume of the highly granular CALICE Silicon-Tungsten electromagnetic calorimeter prototype (SiW ECAL). An algorithm optimised to find interactions in the SiW ECAL at small hadron energies is developed. This allows identifying the interaction point in the calorimeter at an efficiency between 62% and 83% depending on the energy of the primary particle. The unprecedented granularity of the SiW ECAL allows for the distinction between different interaction types. This in turn permits more detailed examinations of hadronic models than was possible with traditional calorimeters. So far, it is possible to disentangle minimum ionising particle (MIP) events, elastic π -nucleus scattering and spallation reactions which lead to the start of a internuclear cascade or which result in a small number of highly ionising particles. Various observables are compared with predictions from hadronic physics lists as contained in the simulation toolkit GEANT4.

Contents

1	Introduction	3
2	The SiW ECAL prototype	3
3	Data samples	4
3.1	Simulation with various GEANT4 physics lists	5
3.2	Event selection	6
4	Investigating the interactions of hadrons	9
4.1	Finding an interaction	10
4.2	Classification of the interactions	13
4.3	Concluding remarks	15
5	Results	16
5.1	Rates of interaction types	16
5.2	Lateral shower extension	17
5.3	Longitudinal profiles	24
5.4	Longitudinal profiles per class of events	24
5.4.1	“FireBall” events	24
5.4.2	“Pointlike” events	25
5.4.3	Longitudinal profiles for “MIP” and “Scattered” events	26
6	Summary, Conclusions and Outlook	27

1 Introduction

The final states of e^+e^- collisions at future lepton colliders comprise high energetic jets containing up to 40 particles. The CALICE (Calorimeter for the Linear Collider Experiment) collaboration conceives and operates prototypes of calorimeters dedicated for the application of particle flow algorithms, *PFA*, foreseen to be employed at a future linear electron-positron collider [1]. Particle flow techniques are promising for achieving a jet energy resolution of 3 – 4% for jets with an energy between 50 GeV and up to more than 500 GeV [2]. The individual hadrons within the jets have typical energies around a few GeV.

The *PFA* approach aims to reconstruct every single particle of the final state of the e^+e^- collision. This goal requires highly segmented calorimeters to disentangle showers created by different particles, i.e. hadrons and photons, within a jet. The optimisation of these particle flow algorithms is supported best if the interactions of hadrons are well modelled and the detector response to hadrons is well understood. In view of this, highly granular calorimeter prototypes provide a unique mean to develop further models of hadronic cascades.

A prototype of a highly granular SiW ECAL [3] was operated in beam test campaigns in 2006 at DESY, 2006 and 2007 at CERN and in 2008 at FNAL. The high granularity of the SiW ECAL permits detailed information to be obtained on interactions of hadrons. This information is turned into global observables describing the shower. Furthermore, it allows for an attempt to classify the different interaction types.

This SiW ECAL is described in the following section. The test beam data and Monte Carlo simulations with the event selections are presented Sect. 3. Interactions of hadrons and the algorithm used for this analysis are presented Sect. 4. Results obtained with the prototype during the test beams at Fermilab with negatively charged pions (π^-) and comparisons with Monte Carlo are discussed Sect. 5. Finally, conclusion and prospects for future studies are given.

2 The SiW ECAL prototype

The SiW ECAL features a sandwich structure comprising 30 layers of silicon (Si) as active material, alternated with tungsten (W) as absorber material.

The active layers are made of Si wafers segmented in $1 \times 1 \text{ cm}^2$ pixels (or pads). As shown in Fig. 1, each wafer consists of a square of 6×6 pixels and each layer is a matrix of 3×3 of these wafers resulting in an active zone of $18 \times 18 \text{ cm}^2$.

The ECAL is divided in 3 modules of 10 layers. The W depth per layer is different in each module increasing from 1.4 mm ($0.4 X_0$) in the first one,

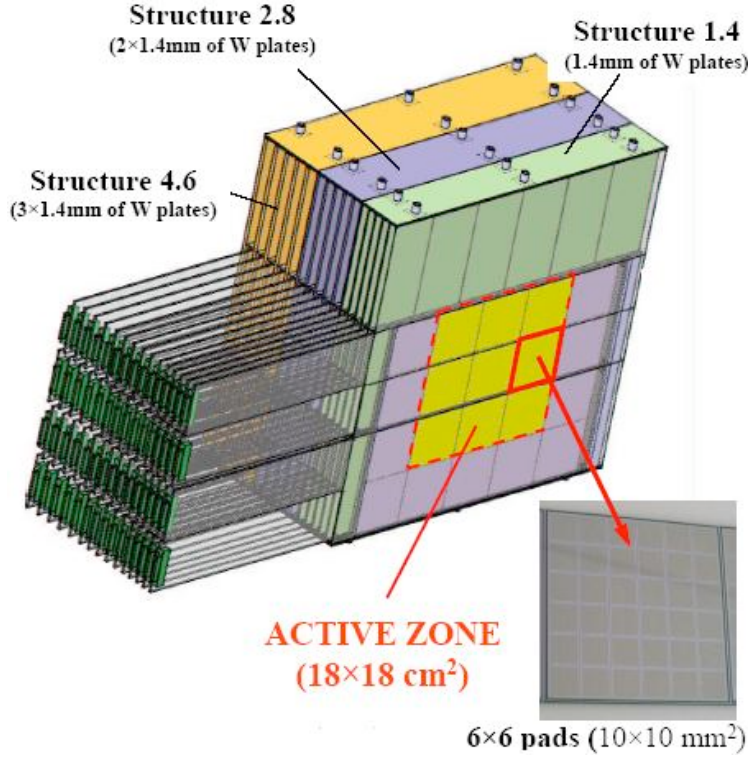


Figure 1: *Schematic view of the SiW ECAL prototype*

to 2.8 mm in the second and 4.2 mm in the last one. This corresponds to $24 X_0$ or $\sim 1 \lambda_I$ which ensures that more than half of the hadrons will have a primary interaction in the ECAL.

3 Data samples

Test beams were conducted in May and July 2008 at the Fermilab Test Beam Facility [5] at FNAL which provides a beam of particles (e^\pm , p , μ^\pm , π^\pm) at energies ranging from 1 GeV to 66 GeV. The ECAL was placed in front of the other CALICE prototypes: an analogue HCAL and a TailCatcher (TCMT) [6]. The beam line was instrumented with a Cherenkov detector used to identify particles, several scintillators used to reject particles and to trigger, and four drift chambers. The scheme of this setup is shown Fig. 2. The coordinate system is right handed with the z axis pointing into the beam direction.

The analysed data in this article comprise runs with primary π^- s. The energies of the primary particles are 2, 4, 6, 8 and 10 GeV. This energy range covers the typical energies of hadrons within a jet. Higher energies were studied in [7] with an overlap at 8 and 10 GeV with the study presented

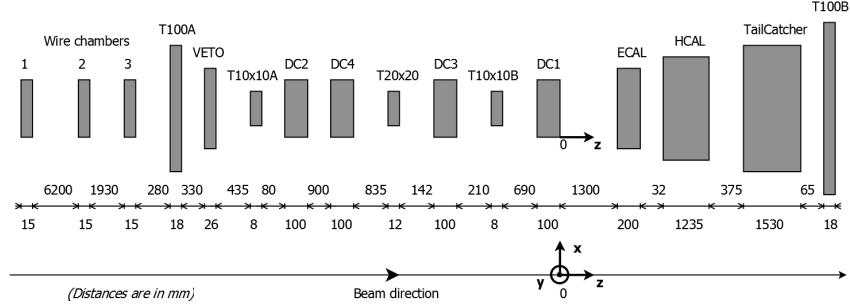


Figure 2: The beam line at FNAL. Distances are in mm.

here.

3.1 Simulation with various GEANT4 physics lists

Due to the complicated nature of hadronic interactions a precise description of hadronic showers in simulations is difficult to achieve. Several models called physics lists are proposed and the high granularity of the SiW ECAL offers unprecedented means to discriminate among them. To compare the test beam data with these models, Monte Carlo simulations were done within the MOKKA framework [8] which provides the geometry interface to GEANT4 [9]. The physics lists used are briefly reviewed here (for more details, see [10] and [11]).

- QGSP_BERT: combines the Bertini model BERT at low energies, making a transition to the Low Energy Parametrized model (LEP) between 9.5 GeV and 9.9 GeV and a further one at energies between 12 and 25 GeV, to the Quark-Gluon-String Precompound model (QGSP). It is used for LHC calorimeters and will be used as reference for this study.
- QGS_BIC: the Binary cascade model (BIC) is used at low energies below 1.2 GeV and for re-scattering of secondaries, then LEP in the intermediate region until 12 GeV and QGS at higher energies.
- QGSP_BIC: uses the BIC model but not for pions for which LEP is used and without re-scattering of secondaries, then LEP below 12 GeV and QGSP above are used.
- LHEP: combines the physics lists LEP below 55 GeV and HEP above 25 GeV, the transition region being thus 25 GeV to 55 GeV.
- FTFP_BERT: uses BERT below 5 GeV and the Fritiof model FTFP above 4 GeV.

The energy ranges in which subcomponents of the introduced models are valid are summarised in Table 1.

E (GeV)	2	4	6	8	10
QGSP_BERT	BERT				BERT + LEP
QGS_BIC	LEP + BIC (secondaries)				
QGSP_BIC	LEP				
FTFP_BERT	BERT		FTFP		

Table 1: *Model used for hadronic interactions depending on the physics list and energy of the interacting particle.*

The CALICE software [12] (v02-00) was used for both data reconstruction and digitisation. This comprises the GEANT4 version 9.3 used for simulations. The starting point of a primary particle is positioned 160 m upstream of the ECAL surface.

3.2 Event selection

Events are triggered by scintillator counters and Cherenkov counters to avoid electron contamination. For further event selection, the following steps are applied as outlined in the following. Where possible, data and simulation are subject to the same selection chain.

- A threshold of 0.6 MIP is chosen to remove noisy hits. Lowering the threshold to 0.4 MIP does not change the results presented in the following study. In particular, finding low energetic interactions which may be subject to fluctuations, is not sensitive at all to this threshold.
- A hit is called isolated if all of the 26 cells in the surrounding cube are empty. Isolated hits are discarded.
- The total number of hits in the ECAL is required to be at least 25 to remove particles which hit the ECAL at a large angle, towards the acceptance limits.
- The barycentres \bar{x} and \bar{y} of the hits are calculated:

$$\bar{x} = \frac{\sum_{hits} x_{hit} E_{hit}}{\sum_{hits} E_{hit}} \text{ and } \bar{y} = \frac{\sum_{hits} y_{hit} E_{hit}}{\sum_{hits} E_{hit}} \quad (1)$$

Requirements are $-50 \text{ mm} < \bar{x} < 50 \text{ mm}$ and $-50 \text{ mm} < \bar{y} < 50 \text{ mm}$ to reduce lateral shower leakage.

- In some events, the wafer in the middle of the bottom part in the 29th layer is showing more than 8 hits whereas no activity is seen around. It is a known problem and the solution adopted here is to remove these events. A fraction of 0.3% of the events suffer also from noise in other layers even after offline corrections. These events are discarded as well.
- Pions are selected using the Cherenkov counters. This selection is refined by counting the number of hits in ECAL, HCAL and TCMT. The Figure 3 shows the number of hits in the SiW ECAL, HCAL and TCMT created by simulated muons. The hit distribution suggest the following cuts to reject muons:

$$N_{\text{ECAL}} < 50, 30 < N_{\text{HCAL}} < 70, 10 < N_{\text{TCMT}} < 35 \quad (2)$$

At 2 GeV, the energy loss of muons being about 1.4 GeV in the HCAL, the conditions need to be slightly changed to match the fewer counts in the TCMT (with $N_{\text{TCMT}} > 5$). The introduced cuts are also indicated in Figure 3. If an event passes these requirements, it is counted as a muon. Using these criteria it is estimated that a pion beam is contaminated with 13% muons at 2 GeV and only 2% at 10 GeV.

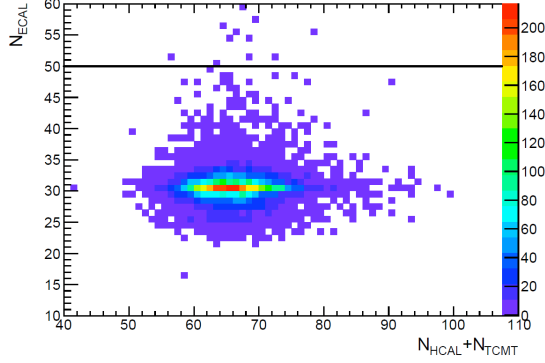
The survival fraction of muons and the rejection fraction of pions after the selection are estimated with a simulation of each 10000 muons and pions. The results are listed in Table 2. Note that the number of rejected pions is largely independent of the physics list where at small energies it is slightly smaller for LEP based models.

E (GeV)	2	4	6	8	10
Remaining muons (%)	2.0	0.4	0.4	0.5	1.0
Rejected pions (%)	8.7	3.0	2.7	2.3	0.9

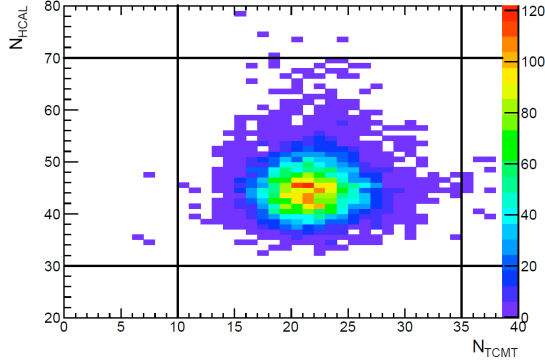
Table 2: *Rates of remaining muons and rejected pions with cuts on muons. The results are obtained by a simulation study.*

From the numbers on muon contamination given before and the entries in the table it can be concluded that the residual muon contamination of the selected sample is well under control. Contamination from electrons is found to be smaller than 1% at 2 GeV and negligible at other energies, according to a Monte Carlo study.

- An algorithm is introduced to identify events in which more than one particle hits the ECAL. These particles can either be genuine multiple particle events caused by beam impurities or products of decays or upstream interactions of primary particles.



(a) Histogram showing the number of hits in the ECAL versus the total number of hits in the HCAL and TCMT for events of 10 GeV muons.



(b) Histogram showing the number of hits in the HCAL versus the number of hits in the TCMT for simulated events of 10 GeV muons.

Figure 3: *Histograms of the number of hits found in the three calorimeters for events with 10 GeV simulated muons. The cuts for muons are then deduced. They are chosen to be energy independent, apart from 2 GeV where a small change is needed.*

The algorithm introduces a distance criterion d_{th} which is calculated from the three dimensional distance between centres of cells carrying a signal. For a given distance $d < d_{th}$ cells are merged into a cluster. Otherwise they seed a new cluster. For this study the algorithm is applied to the first six layers to identify particles entering the SiW ECAL and avoid regions where the signals created by the particles merge due to their interaction in the calorimeter volume. Finally, the algorithm only accepts clusters with more than three hits. The separation power and hence the optimal value of d_{th} depends on the actual cell size but also of the shifts introduced by the staggering of the calorimeter layers.

The value of d_{th} is optimised with the help of a simulation study. For this in each case two muons of 10 GeV from simulated beams incident on the SiW ECAL surface with a width of $\sigma_x = 7.4$ mm in x -direction and of $\sigma_y = 4.5$ mm in y -direction are randomly overlaid. The result is shown in Fig. 4. For a value of $d_{th} = 12$ mm 80% of the muons can be successfully separated. Towards larger values hits from different particles are merged into the same clusters. Therefore the separation power decreases. The sharp drop off towards smaller values of d_{th} is consistent with the lateral cell size of 1×1 cm². In the following the value of $d_{th} = 12$ mm is chosen to identify single particle events.

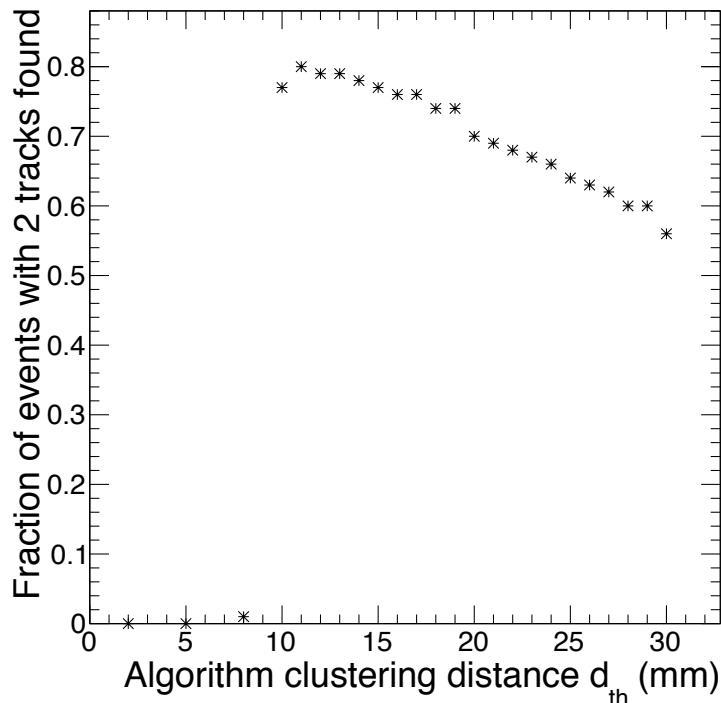


Figure 4: *Efficiency to separate two randomly overlaid 10 GeV muon tracks. It is shown for different separation distances allowed between two hits, below which they are merged into a single cluster.*

The number of selected events is given Table 3.

4 Investigating the interactions of hadrons

The destiny of a primary particle impinging on the SiW ECAL can be twofold. Either the particle passes the ECAL as an ionising particle or it undergoes interactions which lead to the creation of secondary particles.

E (GeV)	2	4	6	8	10
Events	10925	79296	52396	147876	346148

Table 3: Number of events remaining after all the selection criteria are applied to the data.

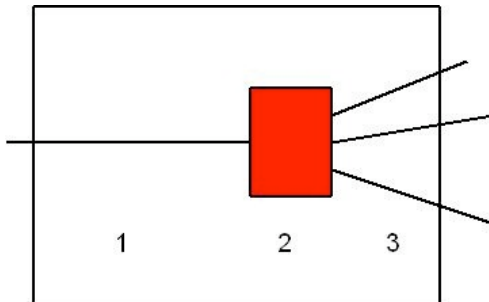


Figure 5: Generic picture of a typical hadronic interaction. 1: a primary track. 2: area of interaction. 3: secondaries emerge from the interaction zone.

The latter case is illustrated in Fig. 5. This figure suggests that each event may be subdivided into three parts. First there is a primary track in the beginning of the ECAL. Second, the interaction occurs. Third, secondaries emerge from the interaction region. Figure 6 shows this behaviour with a test beam event. The first task is thus to identify the interaction layer. The method mentioned above to find tracks entering the ECAL is employed here to define the position and direction of the primary track.

4.1 Finding an interaction

A typical event featuring a large number of secondaries is displayed Fig. 6. The longitudinal profile can be seen in the bottom right histogram. There, upon visual inspection, the deposited energy is significantly increasing in layer 11. This layer is obviously the interaction layer. This is also supported by the lateral view of the event as shown on the other parts of Fig. 6. This condition of interaction can be written:

$$E_i > E_{\text{cut}} \quad \text{and} \quad E_{i+1} > E_{\text{cut}} \quad \text{and} \quad E_{i+2} > E_{\text{cut}} \quad (3)$$

where E_i is the energy deposited in layer i in units of MIP. That is, a cut (E_{cut}) is applied on the deposited energy in each layer. If three consecutive layers have an energy higher than this fixed E_{cut} , the interaction layer is the first of these (layer i).

However this simple cut is not sufficient to find all interactions. Particularly at small hadron energies shower fluctuations are expected to be

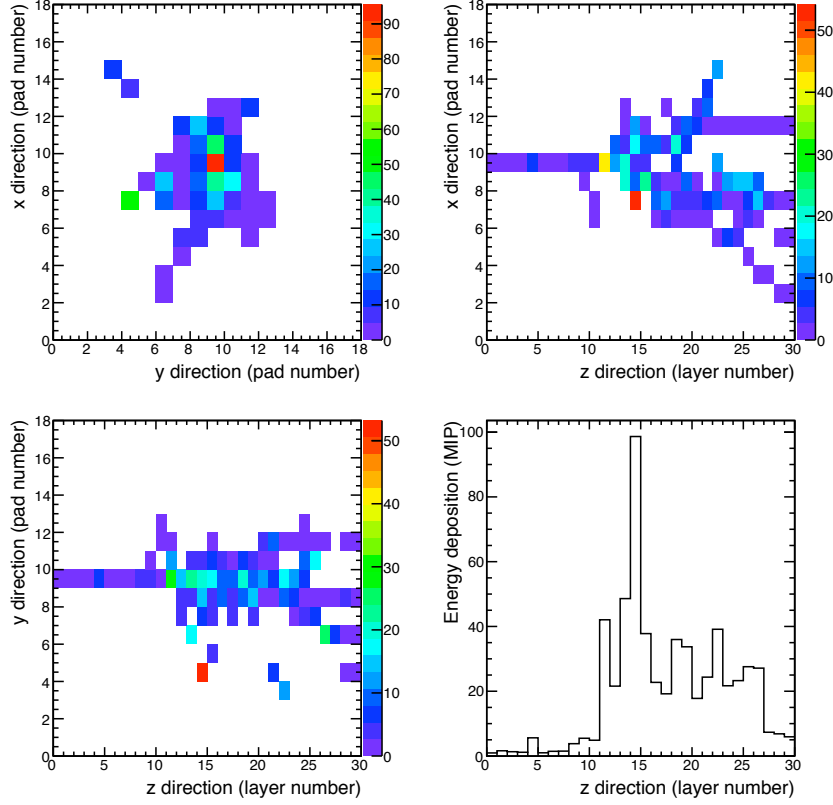


Figure 6: *Display of a hadronic interaction in a test beam event for a pion with an incident kinetic energy of 10 GeV. These are 2D energy weighted profiles of a hadronic interaction in the ECAL: the window on the top left is a projection in the x-y plane of the energy deposited, the one on the top right is the very same projection on the x-z plane and the one on the bottom left is for the y-z plane. The bottom right histogram shows the energy deposition in each layer. Units are cell index in x and y and layer number in z. All start from 0. The energy unit is in MIP.*

strong and interactions may be missed for too high values of E_{cut} . If on the other hand E_{cut} is chosen to be too small, a large fraction of non-interacting events may be accepted. In order to account for fluctuations in the energy deposition, two new variables F and F' are introduced. These are defined as

$$F = \frac{E_i + E_{i+1}}{E_{i-1} + E_{i-2}} \quad \text{and} \quad F' = \frac{E_{i+1} + E_{i+2}}{E_{i-1} + E_{i-2}} \quad (4)$$

They measure a relative increase of energy deposition before and after a given layer i . The fact that two consecutive layers are grouped together makes the variables less sensitive to fluctuations in the energy deposition

when identifying interacting events.

A Monte Carlo study of the variables F and F' is shown in Fig. 7 for 6 GeV pions. Only events with an endpoint inside the ECAL have been considered. The location of the endpoint defines the layer i in Equation 4. In the figure two regions can be distinguished.

- One region with $F \simeq 1$ and $F' \simeq 1$ that represents events in which the primary particles acts MIP-like or in which the relative increase in deposited energy is comparatively small.
- Another region with $F > 1$ or $F' > 1$. These are interacting events.

The discussion before motivates to introduce a new threshold value F_{cut} . If both, F and F' , fulfill the condition

$$F > F_{\text{cut}} \quad \text{and} \quad F' > F_{\text{cut}} \quad (5)$$

the energy deposition is not considered MIP-like anymore. A further study reveals that 10% of the events for which $F > 6$ and $F' > 6$ were rejected by the naive criterion as given by Eq. 3.

The improvement in the recognition of interacting events is confirmed in Table 4. It shows the fraction of interacting events found by using criterion 3 only, with $E_{\text{cut}} = 10$ MIPs and the fraction added when refining the selection using F and F' with $F_{\text{cut}}=4$.

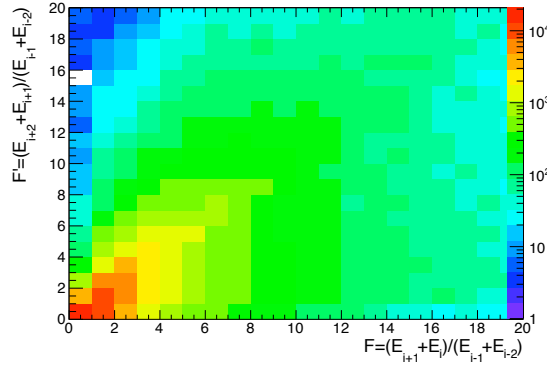


Figure 7: Values of F and F' defined Eq. 5 for simulated 6 GeV pions. The interaction layer given in the Monte Carlo table is set to layer i .

Among the events found using Eq. 5, new topologies leading to a smaller number of secondaries appear. An example is shown in Fig. 8. This event features a strong local increase of energy. This event would have been rejected by Eq. 3. It is recovered by the refinement of the analysis using the introduced variables F and F' . The frequency of these type of events w.r.t. others will be discussed further down.

E (GeV)	f_E (Ecut = 10 MIPs)	$f_{F/E}$ (Fcut = 4)
2	14.3%	26.0%
4	39.6%	20.9%
6	57.9%	15.2%
8	69.1%	11.9%
10	72%	14.9%

Table 4: Table showing the fraction of interacting events found by using criterion 3 (f_E) and those added by using criterion 5 ($f_{F/E}$).

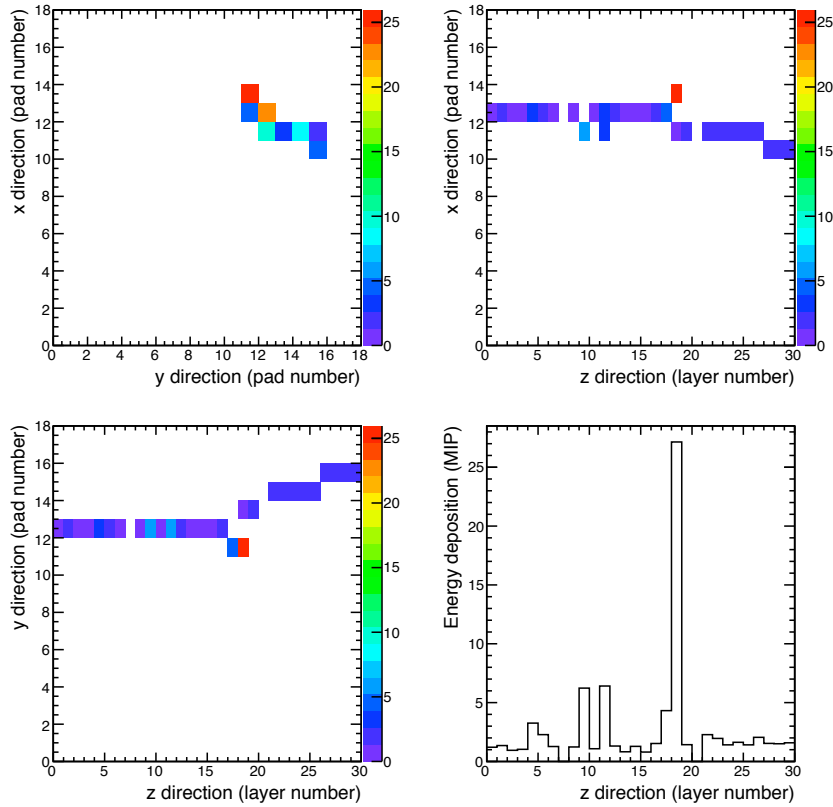


Figure 8: Display of a hadronic interaction in the ECAL for a pion with an incident kinetic energy of 2 GeV (test beam event).

4.2 Classification of the interactions

The algorithm to find interactions in the ECAL has been defined. In a next step, the found interactions are classified.

The interactions passing Eq. 3 will be called “FireBall”. These are typically inelastic hadronic interactions followed by an internuclear cascade. In

the case of other events passing Eq. 5, there are two possibilities.

- The relative increase stops in the fourth layer ($i + 3$) which calls for a local energy deposition. This means:

$$\frac{E_{i+2} + E_{i+3}}{E_{i-1} + E_{i-2}} < F_{\text{cut}} \quad (6)$$

If the relative increase extends only over a couple of layers, it is localised and will be classified as “Pointlike”. This is the case of Fig. 8 where most of the energy released by one secondary particle is localised in a couple of cells. Due to this presence of highly ionising particles, “Pointlike” events can be considered as an opportunity to study details of the short range component of a hadronic shower and are therefore a kind of zoom into that part of the shower. These events might be the result of the evaporation phase of the spallation reaction. The primary particle did not transfer enough energy to the nucleus to release secondaries. Rather the energy is only distributed among the nucleons and excess energy is evaporated in form of invisible neutrons and ionising charged particles. Small energy transfer may also lead to a short truncated internuclear cascade consisting of only a few particles.

As a consequence of this definition, some delta rays will enter this class. The background expected from delta rays is studied with in each case 10k simulated muons at several energies. Since the mass of the muon and of the pion are very close, their behaviour in terms of electromagnetic interactions is very similar. The resulting fraction of events passing the Pointlike criterion is given in Table 5.

E (GeV)	2	4	6	8	10
Rate of delta rays	2.7%	3.7%	4.0%	3.9%	4.1%

Table 5: *Rates of delta rays, estimated to be the fraction of “Pointlike” events found in Monte Carlo simulations of muons.*

- In the second case, the relative increase might continue which means:

$$\frac{E_{i+2} + E_{i+3}}{E_{i-1} + E_{i-2}} > F_{\text{cut}} \quad (7)$$

One has to make sure that this increase is not an artifact caused by a backscattered particle. In the case of a backscattered particle, the relative increase of energy would be caused by the presence of this particle, whose track is several cells away from the primary track. To make sure that the increase is really caused by the start of an hadronic

interaction, one asks for the sum of the energies in the cell of the extrapolated primary track and in the 8 cells around to satisfy, in the reported layer i ($E_{\text{around},i}$):

$$\frac{E_{\text{around},i}}{E_i} > 0.5 \quad (8)$$

This ensures that the increase of energy was caused locally on the path of the track and not by a track away from the initial MIP. This interaction will again be called a “FireBall” because it will be similar to the one Fig. 6 but occurring at lower energies thus depositing less energy. Still it may be regarded as an inelastic reaction.

The remaining events that did not pass those criteria will be considered as non interacting events.

Since the criteria rely on the knowledge of energy deposition in several consecutive layers, some events are discarded: those in which the interaction was found in the three first and three last layers, see Sect. 5.

Discussion of events for which no interaction is identified

Events in which no interaction is found may be separated further into two classes: real MIPs and elastic scattering with negligible nuclear breakup. An example of the latter type is shown in Fig. 9, where a scattered particle is clearly seen. This last kind of event will be classified as a non interacting event for neither Eq. 3 nor Eq. 5 are satisfied. Indeed, the energy deposition is rather small in consecutive layers.

A scattered particle is identified by requiring a lateral distance of at least two pixels between the incoming track and the end of the outgoing track. If the event satisfies this requirement, it will be classified as “Scattered”. If it is not, then it is a real MIP and put in a class called “MIP”. Future studies will cluster the outgoing track and measure the angle with respect to the primary track.

4.3 Concluding remarks

A procedure to find interactions and classify four event types has been presented. With optimised values of E_{cut} and F_{cut} the efficiency to find an interaction is as high as 62% at 2 GeV and up to 83% at 10 GeV. The optimisation of the parameters as well as comparisons with other methods are presented in the appendix of this note.

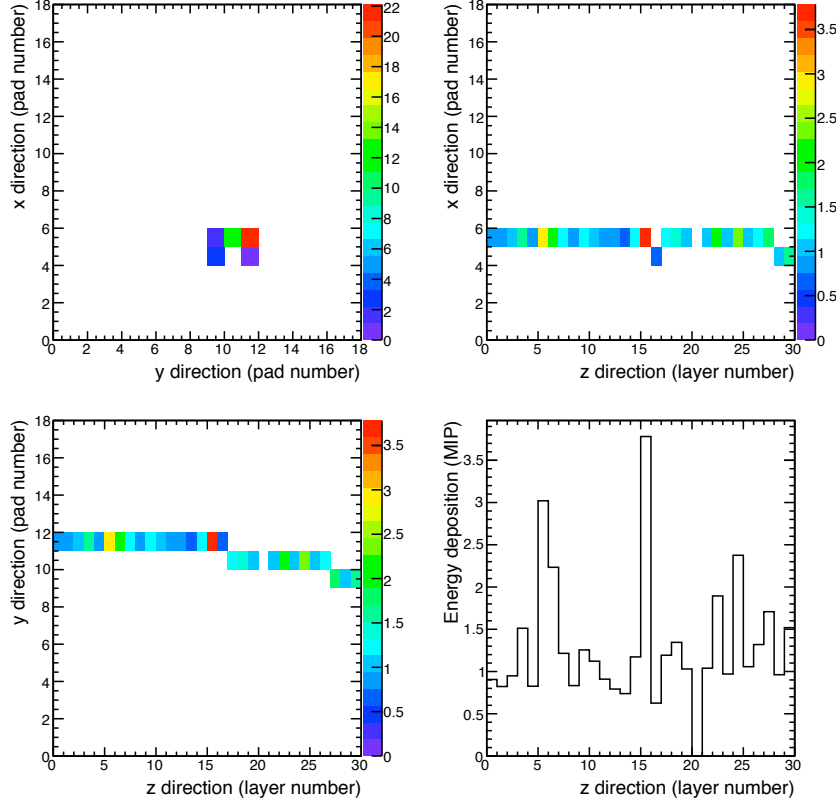


Figure 9: *Display of a measured event at 2 GeV undergoing elastic scattering with negligible nuclear breakup. The event does not fulfill Eq. 3 and Eq. 5 and the track has changed direction in the ECAL. This event is of the type “Scattered”.*

5 Results

The events are classified according to the criteria above and compared with the Monte Carlo predictions by means of their rate, shower radius and longitudinal shower profiles. For a given physics list, the evolution of each profile with respect to the energy is shown. Then detailed comparisons with the various physics lists are made at 2 and 8 GeV.

5.1 Rates of interaction types

The rates of interaction types are defined by the number of events for a given class, divided by the total number of events. Fig. 10 shows the rates of interactions for different energies. The frequency of “FireBall” events in the data is always about 55% nearly independent of the energy. If corrected for the efficiency of finding an interaction, the frequency would increase slightly

towards small energies. The observed dependency is compatible with the behaviour of the π -proton cross section [13] as one of the main underlying scattering processes. The events of type “Pointlike” constitute a complementary fraction of inelastic events. When corrected for δ rays, their frequency is about 4% at small energies and tend towards zero at higher energies. The frequency of both classes for inelastic events is well reproduced by all physics lists which confirms that the total inelastic cross sections are well implemented into the physics lists. This is on the other hand expected since all physics lists implement the underlying π -nucleon scattering at least for the first stage of the intranuclear processes.

For charged pions with kinetic energies above about 1 GeV, the elastic scattering cross section is expected to be suppressed w.r.t. the inelastic part [13]. This suppression is observed in the data and well reproduced by all physics lists. The events of type “Scattered” account for about 5% of the events. The slight increase towards small energies is again broadly compatible with the expected behaviour of the elastic cross section [13].

5.2 Lateral shower extension

A measure of the lateral extension of the final state of the four event types is the shower radius. The shower radius is defined as:

$$\langle r \rangle_E = \sqrt{\sigma_{E,x}^2 + \sigma_{E,y}^2} \quad (9)$$

where e.g.

$$\sigma_{E,x}^2 = \frac{\sum_{\text{hits}} x_{\text{hit}}^2 E_{\text{hit}}}{\sum_{\text{hits}} E_{\text{hit}}} - \left(\frac{\sum_{\text{hits}} x_{\text{hit}} E_{\text{hit}}}{\sum_{\text{hits}} E_{\text{hit}}} \right)^2 \quad (10)$$

and the same for y. For the calculation of the observables, only hits in the interaction layer and all subsequent layers are taken into account.

In order to define in the same way a measure of the radius of non-interacting events, i.e. events where no interaction point could be found, the width $\langle r \rangle_E$ is calculated by summing over all hits in the ECAL.

The profiles shown in Fig. 11 contain each event at energies from 2 GeV up to 10 GeV. Two distinct maxima are visible: a sharp one around 5 mm and a broader one at larger values. The first one is the one expected for MIPs that pass through the ECAL in straight line, the second being the one of the interacting hadrons which contribute to the large radii in the distribution. The data are compared with the prediction from the detector simulation using the QGSP_BERT physics list. For energies smaller than 6 GeV small radii up to about 35 mm are well reproduced by the simulation. Toward higher energies, the transition region between the two maxima is less well described by the simulation. At all energies the simulation does not

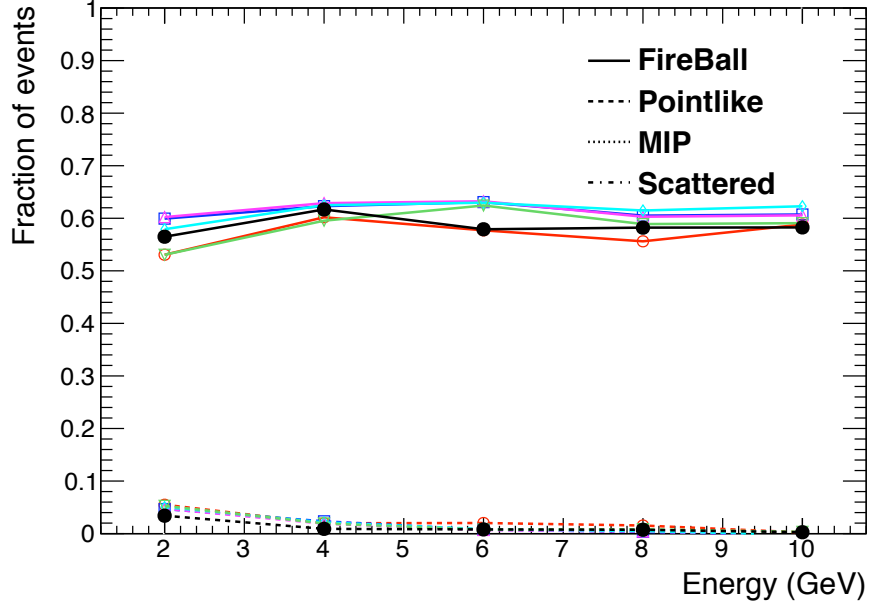
describe the region of large shower radii. The showers seem to be on average broader in the simulation. The description gets however somehow better at the highest energy of 10 GeV.

Using the classification introduced in Section 4, it is possible to separate the contribution of each event type. As can be seen in Fig. 12 at 8 GeV, the hadron peak is seen in the events classified as “FireBall”. The classes of type “Pointlike” and “Scattered” populate the transition region. Finally, the MIP peak is associated with events classified as “MIP” events

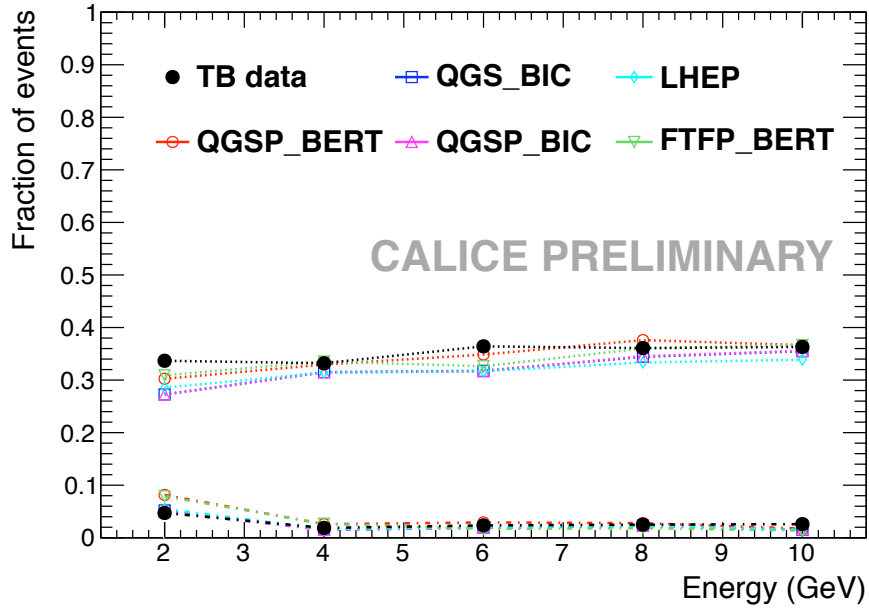
As can be seen in Fig. 12 for “FireBall” events the data undershoot the simulation using QGSP_BERT at large radii which is in agreement with the observation in Figure 11. At small radii the situation is inverted, the data exceed the simulation. An excess in the measured radii w.r.t. the simulation can also be observed for the “Pointlike” events. This observation indicates a deficiency of the simulation to reproduce correctly the topology of weakly populated final states. Note also that naturally there are migrations between the event classes “FireBall” and “Pointlike” and that the migrating events populate the distribution at small radii. This allows for the conclusion that the disagreement in “FireBall” and “Pointlike” at small radii are most presumably of the same origin. The disagreement visible for the “Scattered” and “Pointlike” event types confirms the result that the simulation do not reproduce correctly weakly populated final states. Finally, the disagreement for the “MIP” event type is not expected and currently not completely understood. It might be caused by “Pointlike” events which migrate into the sample of MIP events as the disagreement is most prominent at radii corresponding to the maximum seen in the distribution of the “Pointlike” event types.

The Fig. 13 presents the same comparison as Fig. 12 but now for pions with an energy of 2 GeV. Though conclusions are weakened by the smaller statistics, the distributions at these small energies are broadly reproduced by the simulation. This is particularly true for the events of type “Fireball” with the same tendency for the events of type “Pointlike”. Note, that for pions with an energy 2 GeV the distribution for events of type “MIP” is also better described than for pions with an energy of 8 GeV.

The events of type “Pointlike” and “Scattered” are compared in Fig. 14 with simulations based on the FTFP_BERT physics list. For convenience the comparison with the QGSP_BERT is shown also in the figure. Both distributions are slightly better described by the FTFP_BERT physics list.



(a) Interaction classes



(b) Non interaction classes

Figure 10: Rates of interactions with various physics lists and energies from 2 GeV to 10 GeV. The two graphs in (a) comprise fractions of events with an interaction seen, namely the “FireBall” and “Pointlike” classes, resulting in an interaction layer reported by the algorithm. The two graphs in (b) comprise fractions of events with no interaction seen, the “MIP” and “Scattered” classes. Statistical errors are negligible.

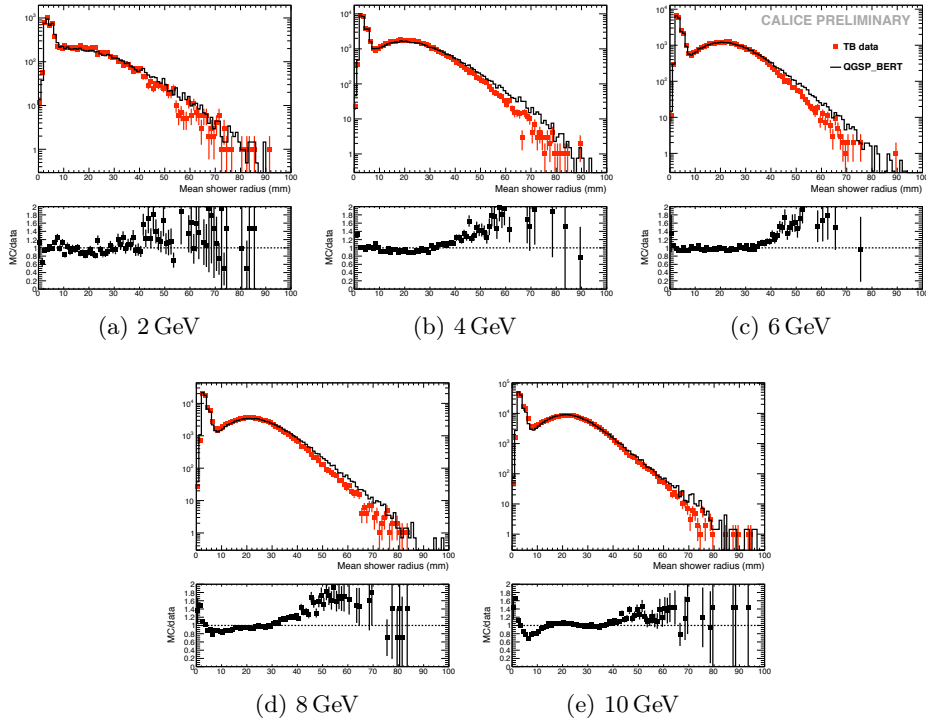


Figure 11: Mean shower radius: the top view features the comparison between test beam data (points with error bars) and QGSP_BERT (solid histograms) for each energy. The bottom view shows the ratio of Monte Carlo simulation and beam test data. The simulation has been normalised to the number of data events for comparison.

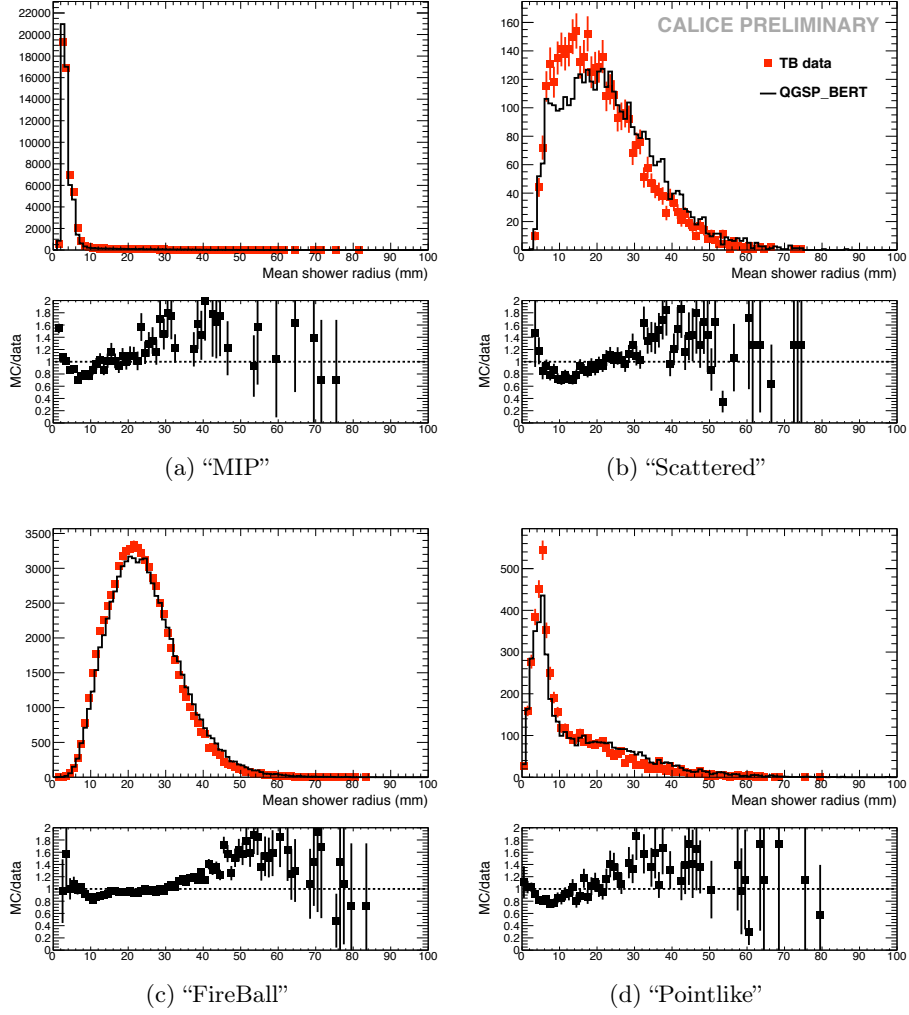


Figure 12: Mean shower radius: the top view features the comparison between test beam data (points with error bars) and QGSP_BERT (solid histograms) at 8 GeV with each class separated. The bottom view shows the ratio of beam test data and simulation. The simulation has been normalised to the number of data events for each class. (a) and (b) show interactions found while (c) and (d) show the non interacting events. The apparently large difference in (b) comes from the larger uncertainties as well as from the physics list itself.

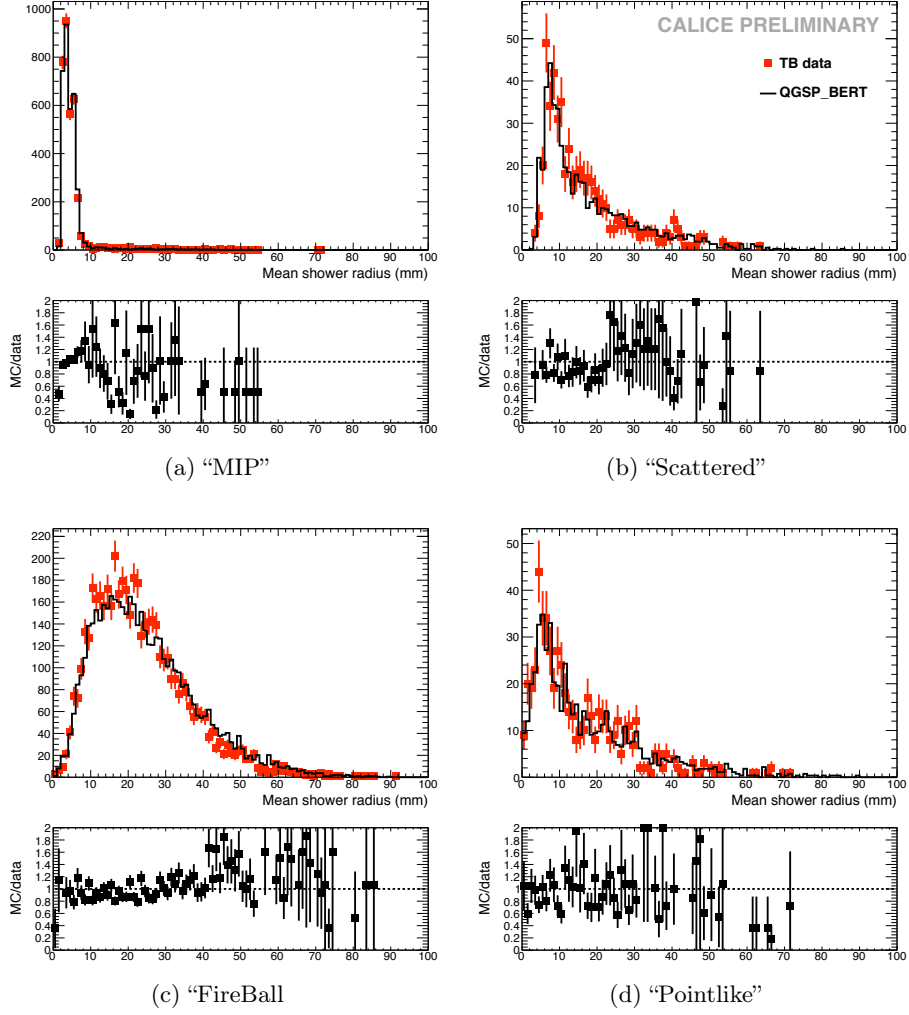


Figure 13: Mean shower radius: the top view features the comparison between test beam data (points with error bars) and QGSP_BERT (solid histograms) at 2 GeV with each class separated, for which the simulation has been normalised to the number of data events. The bottom view shows the ratio simulation and beam test data. The statistics is reduced at 2 GeV and the efficiency of the algorithm is worse. Nevertheless the simulation is still in good agreement with the data. (b) and (d) seem to come from the similar physics processes, as is expected at this small energy.

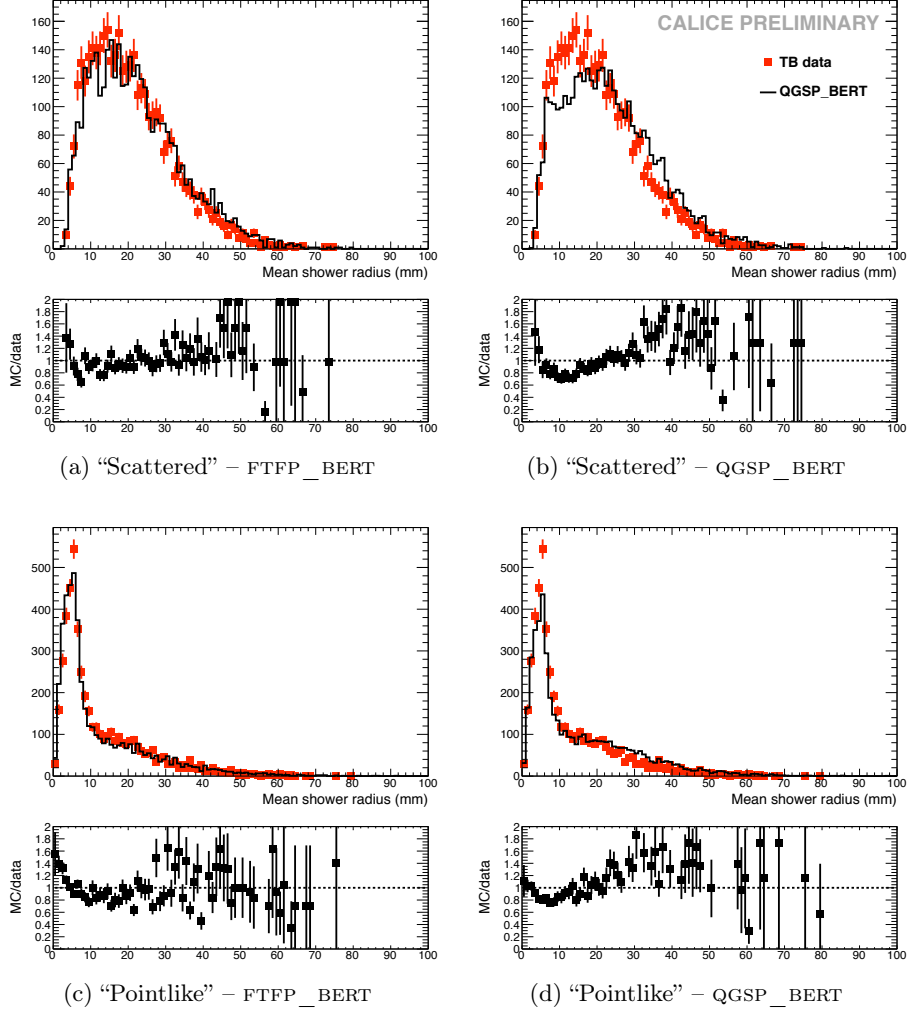


Figure 14: Mean shower radius: the top view features the comparison between test beam data (points with error bars) and QGSP_BERT (solid histogram, left plots) and FTFP_BERT (solid histogram, right plots) at 8 GeV for the “Scattered” class (top plots) and “Pointlike” class (bottom plots), where the simulation has been normalised to the number of data events. The bottom view shows the ratio of simulation and beam test data. The difference seems to come from the physics list itself since FTFP_BERT describes the behaviour correctly. However, statistical uncertainties are to be taken into account and the differences seen with QGSP_BERT may still be acceptable.

5.3 Longitudinal profiles

The longitudinal profiles are defined as introduced in [7]. The longitudinal profile is given as a function of *pseudolayers* in order to account for the different sampling fractions in the ECAL. There is a one to one correspondence between real layers and *pseudolayers* in the first module. On the contrary, each layer in the second module has been subdivided in two *pseudolayers* and layers of the third module have been subdivided into three *pseudolayers*. The energy is then linearly interpolated within the pseudolayers between the energy in the previous layer and the energy in the considered layer. The longitudinal shower profile given in the following histograms starts always from the reconstructed interaction layer. All interactions found between the three first and the three last layers are considered. In case of non-interacting events, the longitudinal profile is calculated from the first detector layer.

5.4 Longitudinal profiles per class of events

For the comparison of data and Monte Carlo in terms of the introduced event classes, those in which an interaction could have been identified, i.e. “FireBall” and “Pointlike” are of primary interest. An example of “MIP” and “Scattered” profile is given for QGSP_BERT, Sect. 5.4.3. For simulated events, the energy contributions for each bin will be decomposed according to the identity of the secondary particle responsible for the energy deposition. In blue, all contributions from electrons and positrons are summed. In green, contributions from protons can be seen. In red, contributions of pions, that is: MIP-like particles are drawn. Finally, the violet histograms are contribution from other particles. The black histogram representing the total is shown for direct comparison with the real data points.

5.4.1 “FireBall” events

In Fig. 15, the “FireBall” events are shown for the Bertini-based models (QGSP_BERT and FTFP_BERT) and LEP based models at 2 GeV. None of the models give a satisfactory description of the data. In Bertini based models, the largest contribution to the profile comes from protons but the energy in the shower maximum is slightly underestimated. In other models using the LEP model, the component labelled “others”, which contains nuclei heavier than hydrogen, is dominant in the first layers and the energy in the shower maximum is overestimated. The models based on the Bertini cascade lead to a slightly better description of the tail of the longitudinal profiles. In this case it looks as if the smaller proton component in case of the LEP physics is responsible for the difference in the simulated distributions.

The Fig. 16 shows the corresponding distributions at 8 GeV. Again no satisfactory description of the entire profile can be achieved. The region around the maximum is better approximated by the QGSP_BERT and FTFP_BERT

physics lists. On the other hand the tails are better described by the QGSP_BIC physics list. Here again the difference between the lists in the contribution from the proton component of the shower is striking.

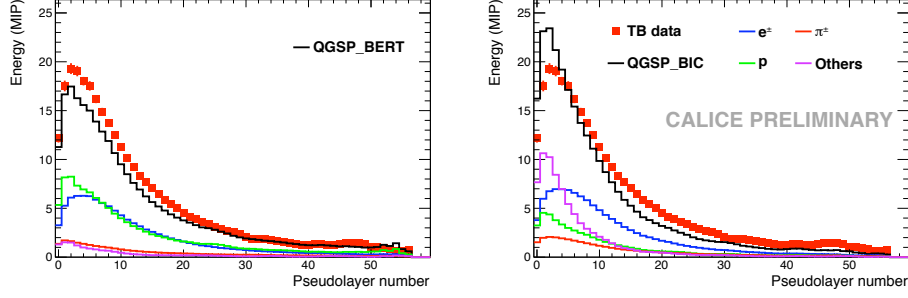


Figure 15: Longitudinal profiles: comparison between test beam data (points with error bars) and QGSP_BERT and QGSP_BIC lists (solid histograms) at 2 GeV for selected “FireBall” events.

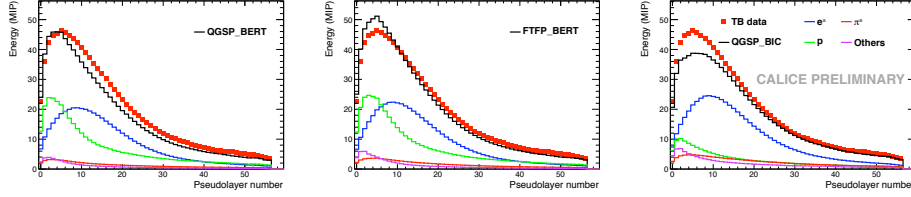


Figure 16: Longitudinal profiles: comparison between test beam data (points with error bars) and QGSP_BERT, FTFP_BERT, QGSP_BIC lists (solid histograms) at 8 GeV for selected “FireBall” events.

5.4.2 “Pointlike” events

The “Pointlike” events as shown Fig. 17 and Fig. 18 feature a strong peak at early layers and then a sharp drop off after 6 pseudolayers. This sharp drop is partially due to the chosen cut scenario but reflects also the short travel distance of highly ionising particles. It is remarkable that the height of the peak is almost the same at both energies. This indicates that at both energies the same mechanism is responsible for these kind of events. As shown in Fig. 17 at 2 GeV this particular class of events favors the BERT physics list which features a dominant proton component. The LEP physics list overshoot the measured spectra. This list realises a large part of the the energy deposition by particles labelled as “others”, namely nuclei with $A > 1$. At 8 GeV, see Fig. 18, the Pointlike events are relatively well modelled by both, the FTFP_BERT and the QGSP_BERT physics list. Differences in the predictions are due to the different energy depositions by heavier nuclei,

again labelled as “others” in Fig. 18. The list QGSP_BIC based on the LEP model fails completely to reproduce the measured spectrum. In contrast to the situation at 2 GeV the energy deposition by heavy nuclei is drastically smaller. On the other hand none of the other components gets enhanced.

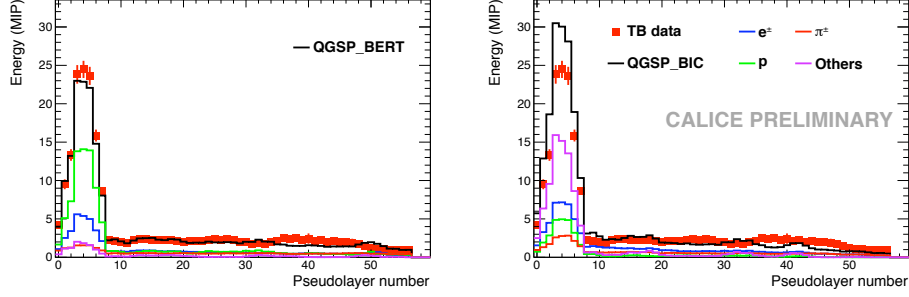


Figure 17: Longitudinal profiles: comparison between test beam data (points with error bars) and QGSP_BERT, QGSP_BIC lists (solid histograms) at 2 GeV for “Pointlike” events.

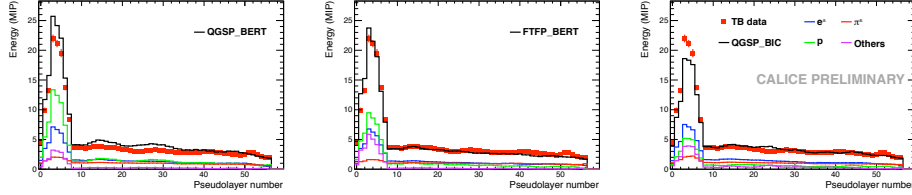


Figure 18: Longitudinal profiles: comparison between test beam data (points with error bars) and QGSP_BERT, FTFP_BERT, QGSP_BIC lists (solid histograms) at 8 GeV for “Pointlike” events.

5.4.3 Longitudinal profiles for “MIP” and “Scattered” events

For completeness, the longitudinal profiles obtained for events with no interaction found is shown Fig. 19. The physics list is QGSP_BERT.

For the “MIP” part, no discrepancy is seen but an increase in the energy deposition in the last *pseudolayers*. Having a look at the proton content which is representative for the start of showers as seen Sect 5.4.1, one can infer that this comes from events where the interaction was not found and thus enter the “MIP” class. Since the interaction occurs in the last layers of the ECAL, it is clearly due to the algorithm which cannot find interactions in the two last layers.

A particular alternated structure is seen for the electromagnetic subcomponent. This is the odd-even layer effect due to the alternated structure of the layers, see also [4].

The longitudinal profile for the “Scattered” events gather events missed by the cuts but still contain information on interactions with low energy deposition. Due to the loose energy cut at 2 GeV of $E_{\text{cut}} = 3$ MIPs, it is not likely that interacting events will be a large part of this class, see Fig. 19 (b), apart from later interactions which needed a further cut to remove noise. On the contrary, since the cut at 8 GeV is tighter ($E_{\text{cut}} = 8$ MIPs), more interacting events that have not been found will be left. The deposited energy increases with increasing detector depth. This is plausible since an interaction at larger depth is more likely to be missed than an interaction occurring in early layers.

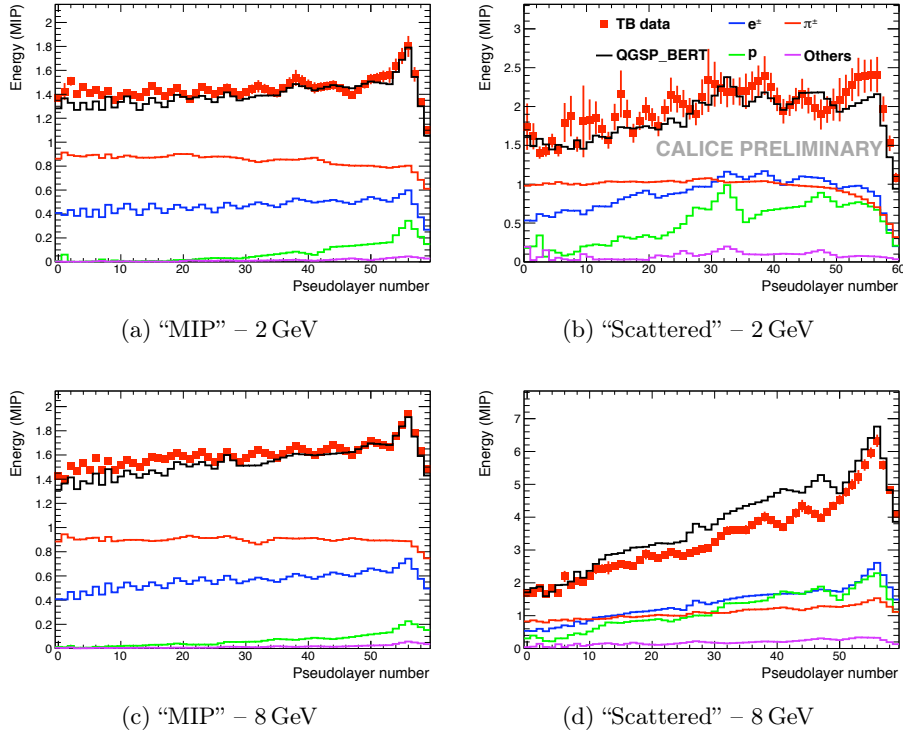


Figure 19: *Longitudinal profiles: comparison between test beam data (points with error bars) and QGSP_BERT (solid histogram) at 2 GeV and 8 GeV for the two classes with no interaction found.*

6 Summary, Conclusions and Outlook

A deep understanding of hadronic showers is of general interest and naturally supports the future development of particle flow algorithms for detectors at a future linear collider. This study demonstrates the large potential of the CALICE SiW ECAL to obtain a detailed image of hadronic cascades.

The data obtained in test beams with pions of an energy between 2-10 GeV are compared with Monte Carlo predictions employing different physics lists as contained in the simulation toolkit GEANT4. The start of the hadronic shower in the interacting case can be reconstructed to an accuracy of better than two layers with an efficiency between 62% and 83% depending on the energy of the primary particle. A classification of the reaction which the pions undergo in the calorimeter volume is realised. Currently, it is possible to distinguish between MIP like events, elastic scattering events, and spallation reactions which lead either to the start of an internuclear cascade or which result in a small number of highly ionising particles. The quality of the description of the data varies with the energy of the primary particle and the chosen physics list. None of the chosen physics lists can describe the entire set of data. Models based on the Bertini cascade, i.e. QGSP_BERT and FTFP_BERT tend to be closer to the data than e.g. the LEP based physics lists.

The next step will be to classify inelastic reactions in terms of shower topology. This comprises the determination of size and energy density of the interaction region as well as the measurements of tracks emerging from the interaction region. These steps will exploit further particularly the lateral granularity of the ECAL. They may form a solid base for the development or the improvement of particle flow algorithms.

References

- [1] J.C. Brient and H. Videau,
The calorimetry at the future e^+e^- linear collider.
eConf C010630 E3047 2001, LLR-02-001, arXiv:hep-ex/0202004v1.
- [2] M.A. Thomson,
Particle Flow Calorimetry and the PandoraPFA Algorithm.
Nucl. Instrum. Meth. A **611** (2009) 25, arXiv:0907.3577v1 [physics.ins-det].
- [3] The CALICE collaboration, J. Repond *et al.*,
Design and Electronics Commissioning of the Physics Prototype of a Si-W Electromagnetic Calorimeter for the International Linear Collider,
JINST **3**, P08001 (2008), arXiv:0805.4833v1 [physics.ins-det].
- [4] The CALICE collaboration, C. Adloff *et al.*,
Response of the CALICE Si-W Electromagnetic Calorimeter Physics Prototype to Electrons,
Nucl. Instrum. Meth. A **608**, 372 (2009), arXiv:0811.2354 [physics.ins-det].
- [5] Fermilab Test Beam Facility, <http://www-ppd.fnal.gov/MTBF-w>

- [6] The CALICE Collaboration, C. Adloff *et al.*,
Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype,
JINST **5**, P05004 (2010), arXiv:1003.2662v1 [physics.ins-det].
- [7] The CALICE Collaboration, C. Adloff *et al.*,
Study of the interactions of pions in the CALICE silicon-tungsten calorimeter prototype,
JINST **5**, P05007 (2010), arXiv:1004.4996 [physics.ins-det].
- [8] Mokka webpage, <http://polzope.in2p3.fr:8081/MOKKA>
- [9] Geant4 webpage, <http://geant4.web.cern.ch/geant4>
- [10] Reference physics lists in Geant4, http://geant4.web.cern.ch/geant4/support/proc_mod_catalog/physics_lists/referencePL.shtml
- [11] V. Uzhinsky *et al.*,
GEANT4 simulation of hadronic interactions at 8–10 GeV/c: response to the HARP-CDP group.
Eur. Phys. J. C **61**, 237 (2009).
- [12] CALICE Software webpage, <https://svnsrv.desy.de/viewvc/calice>
- [13] K. Nakamura *et al.* (Particle Data Group),
J. Phys. G **37**, 075021 (2010).
- [14] R. Wigmans, *Calorimetry – Energy Measurement In Particle Physics*.
Oxford Science Publications.

Appendix - Optimisation of the selection criteria

Monte Carlo simulations are employed to optimise the parameters E_{cut} and F_{cut} . The interaction point of the incoming hadron is defined by the known endpoint of its trajectory. With appropriate parameters, the algorithm should return a reconstructed interaction layer closely to the true one. The parameters E_{cut} and F_{cut} are optimised using the following quantities on Monte Carlo samples of pions at different energies.

Interacting and non interacting events

An interacting event is defined in the following way. First, only events with an interaction point inside the ECAL are considered. For these events, the average energy per hit is calculated for each layer:

$$e_{\text{layer}} = \frac{E_{\text{layer}}}{N_{\text{layer}}} \quad (11)$$

where E_{layer} denotes the energy in the layer, and N_{layer} denotes the corresponding number of hits.

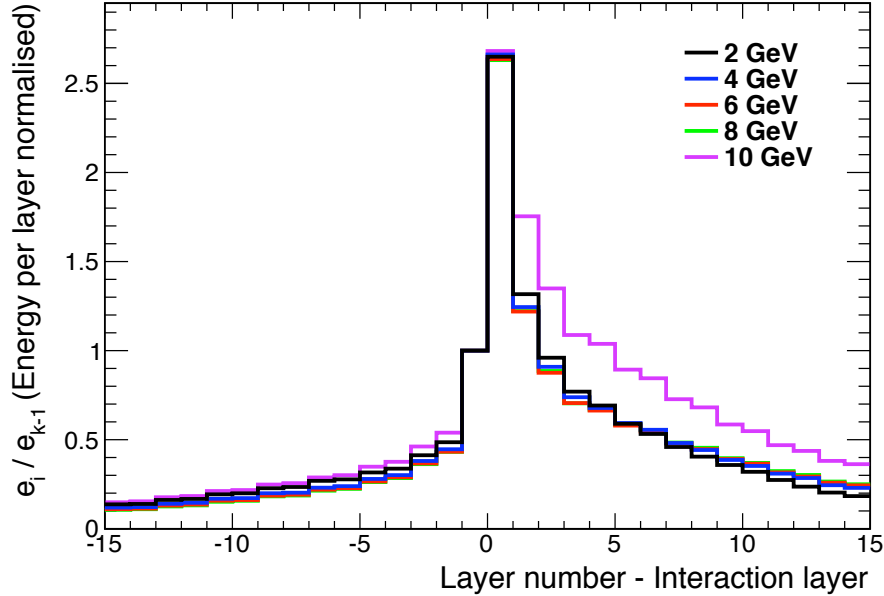


Figure 20: The averaged energy per hit in each layer for all energies, divided by the energy per hit in the last layer before interaction ($k - 1$, taken from the Monte Carlo record). It is centered around the interaction layer (thus layer 0) and each bin is normalised by its number of counts. (Simulation study)

Second, deducing from the endpoint the interaction layer k , the corresponding distribution is shown in Fig. 20 for all energies, featuring the energy per hit for each layer e_i divided by the energy per hit in the last layer before interaction, e_{k-1} . This plot suggests that interacting events in the Monte Carlo be selected using:

$$e_k \geq 1.2 e_{k-1} \quad (12)$$

It has the clear physical meaning that the energy deposited after the interaction should be higher than before. This reflects the idea that the parameters E_{cut} and F_{cut} will be optimised to find events of the type “FireBall” and “Pointlike” i.e. where an interaction occurred.

Non interacting events are complementary: those where the particle did not interact in the ECAL volume or those rejected by Eq. 12. This definition of non interacting events not only includes MIPs but also events with elastic scattering. The events of the “MIP” and “Scattered” classes should thus belong to the non interacting events.

The cuts are optimised according to the following criteria:

- **Standard deviation**

The difference between the reconstructed layer and the true interaction layer is minimal. Thus, the interesting quantity is the standard deviation of this difference. It is obtained by fitting with a gaussian the distribution of *reconstructed layer* - *true layer* in a range of ± 5 layers around 0. It is naturally calculated with interacting events only.

$$\text{Standard deviation} = \sigma_{\text{gaussian fit}}(\text{reconstructed layer} - \text{true layer}) \quad (13)$$

- **Interaction fraction**

The interaction fraction I_f is defined by the rate of events where an interaction was found by the algorithm among interacting events:

$$I_f = \frac{\text{Number of interactions found}}{\text{Number of interacting events}} \quad (14)$$

It is to be maximal.

- **Purity with non interacting events**

The last test variable measures the purity P of the algorithm. The purity is defined by the number of non interacting events identified by the algorithm among the number of non interacting events, i.e.

$$P = \frac{\text{Number of non interacting events returned}}{\text{Number of non interacting events}} \quad (15)$$

Optimisation of Ecut and Fcut

In a first step Ecut is varied between 1 and 20 in steps of 1 while Fcut is kept constant at 6 MIPS. The evolution of the standard deviation and the interaction fraction for pions with an energy of 10 GeV is shown in Fig. 21(a). For small values of Ecut the algorithm interprets already small energy fluctuations as interactions. Due to this arbitrariness of the identification the standard deviation is naturally very large and the interaction fraction is close to 1. Towards larger values of Ecut the algorithm gets more accurate in identifying the interaction. This leads to a decrease in the standard deviation. At the same time the interaction fraction remains close to 100%. If Ecut is further increased the algorithm tends to return layers in which the shower is already developed which explains the shallow minimum in Fig. 21(left).

For the following optimisation process the absolute minimum of the standard deviation is defined as σ_{min} and the maximum of the interaction fraction is defined as $I_{f,max}$, with its corresponding standard deviation σ_{max} which defines $\Delta\sigma = \sigma_{max} - \sigma_{min}$.

To find an optimal value of Ecut, limits on its standard deviation σ and interaction fraction I_f are gradually increased and decreased by using:

$$\begin{aligned}\sigma(n) &= \sigma_{min} + 0.01n \cdot \Delta\sigma \\ I_f(n) &= I_{f,max} - 0.01n\end{aligned}\tag{16}$$

such that

$$\sigma < \sigma(n) \text{ and } I_f > I_f(n)\tag{17}$$

The smallest value of n that permits to find a value of Ecut defines this optimum. In case of several values of Ecut found, the smallest is kept, in favour of a larger interaction fraction.

Figure 21(right) shows the evolution of the standard deviation and the interaction fraction for pions with an energy of 2 GeV. At small energies the distribution does not transit a minimum as the energy of the secondaries is comparatively small and subject to fluctuations. Still, the same algorithm as developed for larger energies can be used to find the optimal value of Ecut. Note, that for small energies the interaction fraction never reaches 100% which underlines again the need for a second observable for a proper identification of the interaction layer.

The introduced value of Fcut is optimised with the help of the purity as defined in Eq. 15 and the interaction fraction. For a constant Ecut of 10 MIPS, Fcut is varied in steps of 0.5 between 1 and 10. The evolution of the purity for pions with an energy of 10 GeV and 2 GeV is shown in Fig. 22. For small values of Fcut, a lot of interactions are found which gives the largest interaction fraction, but some fake interactions due to fluctuations are found too, associated with a worse purity. When Fcut is increased, the rate of interactions found decreases but the fake interactions tend to decrease as well

thus increasing the purity up to the maximal value P_{max} . When choosing the optimal value of F_{cut} the emphasis is put on maintaining a high purity. The optimal value of F_{cut} is given by the first value for which $P > 0.95 \cdot P_{max}$.

As shown in Figs. 21 and 22 the optimisation is carried out for the two physics lists QGSP_BERT and QGS_BIC. While the actual values of interaction fraction, standard deviation and purity are different, the resulting optimal values are independent of the physics list. This indicates the general applicability of the introduced algorithm.

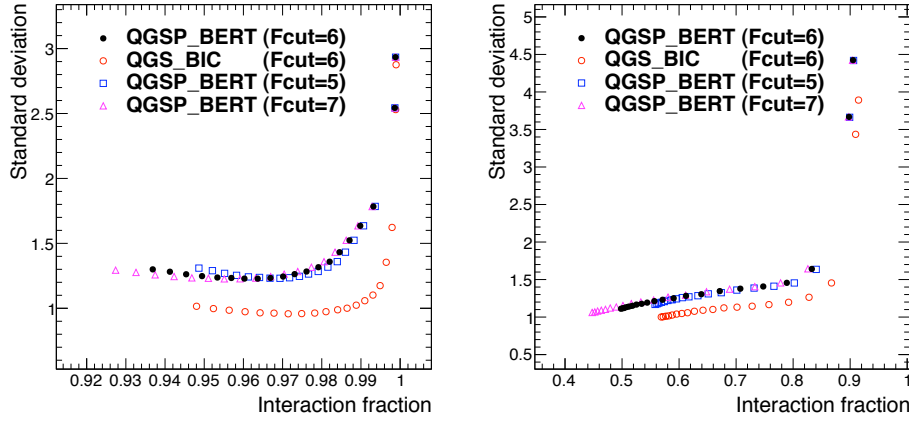


Figure 21: Standard deviation (in units of layer) versus the interaction fraction for pions with and energy of 10 GeV (left) and 2 GeV (right). Results are shown for a fixed $F_{cut} = 6$. The full black dots are for QGSP_BERT and the open red ones for QGS_BIC. The squares represent a change in F_{cut} of $+1$ (full violet) and -1 (open blue) with QGSP_BERT. The energy cut (E_{cut}) is varied from 1 to 20 MIPs by steps of 1 MIP.

Results of optimisation and efficiency of finding interactions

The Table 6 shows the optimal values of E_{cut} and F_{cut} for different energies. The values for E_{cut} are energy dependent as expected while the optimal values for F_{cut} do not vary much with energy. In the following the values for E_{cut} are used as given in the table while $F_{cut}=6$ is chosen for all considered energies. With these values for E_{cut} and F_{cut} , the efficiency η over all interacting events to find the interaction layer in ± 1 layer around the true layer and ± 2 layers around is computed.

The result is given in Table 7. It shows that the efficiency is always larger than 62% where as expected higher efficiencies are obtained for the higher energies. The results for the wider range can be compared with the “3 out of 4” method, i.e. η_{3-4} with a cut at 10 MIPs like in [7]. All samples used are made with QGSP_BERT. It demonstrates that towards small energies the

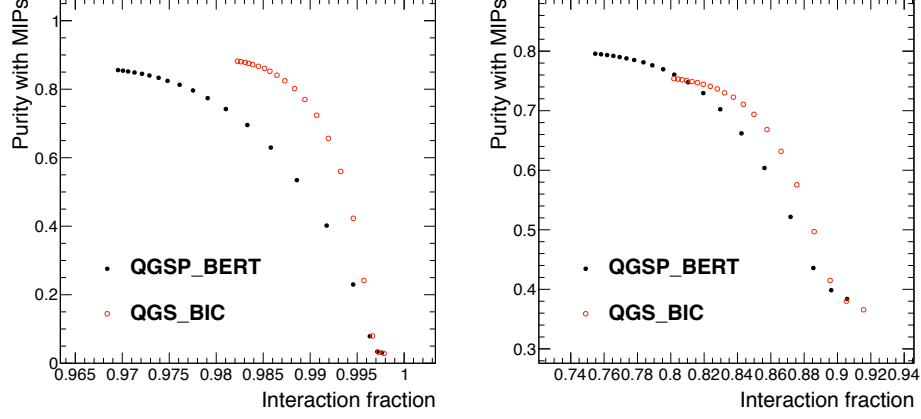


Figure 22: Graph showing purity versus the interaction fraction for primary pions with an energy of 10 GeV (left) and 2 GeV (right). The value of E_{cut} has been fixed to $E_{\text{cut}} = 10$ MIPs for pions with an energy of 10 GeV and to $E_{\text{cut}} = 3$ MIPs for pions with an energy of 2 GeV.

cut scenario presented in this article results in significantly higher efficiencies compared with the simpler method which gets sufficient as the energy increases. The “3 out of 4” method is indeed sufficient at higher energies where most interactions are of the “FireBall” class while at smaller energies, the relative importance of the other classes defined becomes higher.

List E (GeV)	QGSP_BERT		QGS_BIC	
	Ecut (MIP)	Fcut	Ecut (MIP)	Fcut
2	3	5	3	5
4	4	6	4	5.5
6	7	6	7	6
8	9	6	7	6
10	8	6.5	7	6

Table 6: Cuts used at each energy. They are almost energy independent from 6 to 10 GeV.

E (GeV)	$\eta_{\pm 1}$	$\eta_{\pm 2}$	$\eta_{3-4, \pm 2}$
2	0.54	0.62	0.22
4	0.58	0.67	0.51
6	0.62	0.72	0.64
8	0.64	0.75	0.69
10	0.74	0.83	0.78

Table 7: *Efficiency η to find the interaction at each considered energy with the algorithm within ± 1 layer, ± 2 layers, compared with the “3 out of 4” method [7], i.e. η_{3-4} . Only interacting events are considered.*